

УДК 004.032.26

ПОДБОР АРХИТЕКТУРЫ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ РУКОПИСНЫХ ЦИФР

Кузенная А. М., Солдатов О. П.

Самарский национальный исследовательский университет
имени академика С. П. Королёва, г. Самара

В современном мире все чаще приходится слышать о распознавании изображений и применении к этой задаче компьютерных технологий. Один из классических способов решения этой задачи – использование сверточных нейронных сетей. Структура классической свёрточной сети представлена на рисунке 1. Свёрточная сеть состоит из входного и выходного слоя и скрытых слоёв свёртки и подвыборки, однако вопрос о количестве скрытых слоёв в свёрточных сетях остается открытым, то есть остается открытым вопрос об оптимальной архитектуре такой сети, при которой погрешность распознавания будет минимальна.

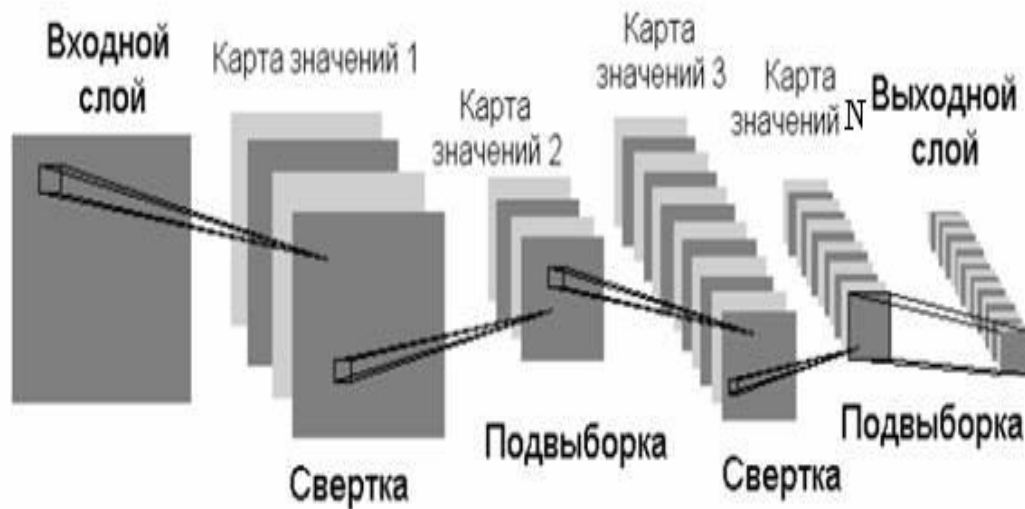


Рис. 1. Архитектура сверточной нейронной сети

Целью работы является такой подбор архитектуры и параметров сети, которые обеспечат минимальную погрешность распознавания тестового подмножества данных, не участвовавших в обучении. Критерием правильности окончательных результатов является погрешность обобщения, вычисленная по тестовой выборке.

Чтобы ограничить полный перебор всех возможных архитектур, обратимся к мере Вапника-Червоненкиса:

$$2 \left\lceil \frac{K}{2} \right\rceil N \leq h \leq 2M(1 + \lg L), [1]$$

где h – сама мера Вапника-Червоненкиса,
 K – количество нейронов в скрытых слоях,
 N – размерность входного вектора,
 M – общее количество весов сети,
 L – общее количество нейронов сети.

Мера Вапника-Червоненкиса (VC-измерение) отражает уровень сложности сети и тесно связана с количеством содержащихся в ней весов. Чем больше число различных весов, тем больше сложность сети и, соответственно, значение VC-измерения. Определив границы нахождения меры Вапника-Червоненкиса для каждой из

рассматриваемых сетей, установим, входит ли каждая из рассматриваемых архитектур в доверительный интервал, заданный допущением.

Со статистической точки зрения, погрешность обобщения E зависит от уровня погрешности обучения H и от доверительного интервала и характеризуется отношением:

$$E \leq H + \varepsilon, [1]$$

где $\varepsilon = (\frac{P}{h}, H)$ – доверительный интервал, P – объем обучающей выборки, h – мера Вапника-Червоненкиса.

Если мера для архитектуры сети входит в доверительный интервал и прогнозируемая погрешность обобщения не ниже заданной, такая архитектура может быть использована для распознавания рукописных цифр. С помощью тестирования всех подобных сетей можно подобрать оптимальную архитектуру для решения данной задачи.

Тестирование и обучение свёрточных сетей для простоты при этом проводится с помощью библиотеки Lasagne языка Python [2]. Эта библиотека позволяет осуществлять подбор архитектуры сети, не прибегая к громоздким многострочным алгоритмам, при этом внутри реализации библиотеки имеется сложная оптимизация, ускоряющая вычисления за счет распараллеливания вычислений между ядрами центрального и графического процессоров. Библиотека позволяет построить сеть любой архитектуры за минимальное человеческое и машинное время.

Для поиска оптимальной архитектуры проводится перебор из всех подходящих сетей по нескольким параметрам: количество слоев, их состав, размер и порядок. Затем на наилучшей найденной архитектуре находятся лучшие функция активации и способ инициализации весов сети.

Для проведения вычислительных экспериментов была использована база рукописных цифр MNIST [3,4], содержащая 60 000 обучающих и 10 000 тестовых изображений в размеченной базе. Проведение анализа распознавания рукописных цифр в зависимости от архитектуры свёрточной нейронной сети показало, что использование меры Вапника-Червоненкиса эффективно справляется с отсечением малоэффективных архитектур, и может быть использована в разработке более сложных систем.

Библиографический список

1. Осовский, С. Нейронные сети для обработки информации [Текст] / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.: ил.
2. Lasagne, Python library – <https://lasagne.readthedocs.io>.
3. LeCun, Y. The MNIST database of handwritten digits –<http://yann.lecun.com/exdb/mnist>.
4. Stanford. The MnistHelper functions – http://ufldl.stanford.edu/wiki/index.php/Using_the_MNIST_Dataset.